AD

Technical Report AREAW-TR-08014

REQUIRED REVISIONS TO CLASSICAL ELECTROMAGNETISM

Paul J. Cote
Mark Johnson

October 2008

ARMAMENT RESEARCH, DEVELOPMENT AND ENGINEERING CENTER
Weapons & Software Engineering Center
Benét Laboratories

Approved for public release; distribution is unlimited.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
15-10-2008	FINAL	

4. TITLE AND SUBTITLE
Required Revisions to Classical Electromagnetism

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Paul J Cote & Mark Johnson

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army ARDEC
Benet Laboratories, RDAR-WSB
Watervliet, NY 12189-4000

8. PERFORMING ORGANIZATION REPORT NUMBER
AREAW-TR-08014

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army ARDEC
Benet Laboratories, RDAR-WSB
Watervliet, NY 12189-4000

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSORING/MONITORING AGENCY REPORT NUMBER

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
It is shown that the generally accepted set of Maxwell's equations is incomplete and an additional law pertaining to the divergence of the induced electric field is required. A major implication is that standard derivations of the wave equations given in the literature are invalid. It is also shown that the standard electromagnetic gauge is fundamentally flawed. The scalar and vector potentials can be addressed without the standard gauge concept, which simplifies the standard formalism.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**	U	22	Mark Johnson
U/U	U	U			**19b. TELEPONE NUMBER (*Include area code*)** 518-266-4986

INSTRUCTIONS FOR COMPLETING SF 298

1. REPORT DATE. Full publication date, including day, month, if available. Must cite at lest the year and be Year 2000 compliant, e.g., 30-06-1998; xx-08-1998; xx-xx-1998.

2. REPORT TYPE. State the type of report, such as final, technical, interim, memorandum, master's thesis, progress, quarterly, research, special, group study, etc.

3. DATES COVERED. Indicate the time during which the work was performed and the report was written, e.g., Jun 1997 - Jun 1998; 1-10 Jun 1996; May - Nov 1998; Nov 1998.

4. TITLE. Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

5a. CONTRACT NUMBER. Enter all contract numbers as they appear in the report, e.g. F33615-86-C-5169.

5b. GRANT NUMBER. Enter all grant numbers as they appear in the report, e.g. 1F665702D1257.

5c. PROGRAM ELEMENT NUMBER. Enter all program element numbers as they appear in the report, e.g. AFOSR-82-1234.

5d. PROJECT NUMBER. Enter al project numbers as they appear in the report, e.g. 1F665702D1257; ILIR.

5e. TASK NUMBER. Enter all task numbers as they appear in the report, e.g. 05; RF0330201; T4112.

5f. WORK UNIT NUMBER. Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

6. AUTHOR(S). Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES). Self-explanatory.

8. PERFORMING ORGANIZATION REPORT NUMBER. Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

9. SPONSORING/MONITORS AGENCY NAME(S) AND ADDRESS(ES). Enter the name and address of the organization(s) financially responsible for and monitoring the work.

10. SPONSOR/MONITOR'S ACRONYM(S). Enter, if available, e.g. BRL, ARDEC, NADC.

11. SPONSOR/MONITOR'S REPORT NUMBER(S). Enter report number as assigned by the sponsoring/ monitoring agency, if available, e.g. BRL-TR-829; -215.

12. DISTRIBUTION/AVAILABILITY STATEMENT. Use agency-mandated availability statements to indicate the public availability or distribution limitations of the report. If additional limitations/restrictions or special markings are indicated, follow agency authorization procedures, e.g. RD/FRD, PROPIN, ITAR, etc. Include copyright information.

13. SUPPLEMENTARY NOTES. Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

14. ABSTRACT. A brief (approximately 200 words) factual summary of the most significant information.

15. SUBJECT TERMS. Key words or phrases identifying major concepts in the report.

16. SECURITY CLASSIFICATION. Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

17. LIMITATION OF ABSTRACT. This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.

Abstract

It is shown that the generally accepted set of Maxwell's equations is incomplete and an additional law pertaining to the divergence of the induced electric field is required. A major implication is that standard derivations of the wave equations given in the literature are invalid. It is also shown that the standard electromagnetic gauge is fundamentally flawed. The scalar and vector potentials can be addressed without the standard gauge concept, which simplifies the standard formalism.

PACS: 41., 41.20.Cv, 41.20.Gz, 41.20.Jb

Contents

Abstract i
Introduction 1
The Missing Maxwell Equation 3
Vector Potential 5
Alternative Approach to the Wave Equations 7
Reconciling with Special Relativity 13
Summary 14
References 15
Appendix 16

Introduction

A review of the basic derivations leading to the standard electromagnetic wave equations reveals a number of errors. These errors and the revisions required to correct the basic formalism are the subject of this paper. The present section lists the basic Maxwell's equations and provides an outline of the standard wave equation derivations for later reference. Subscripts and superscripts are used in an attempt to provide precision to the definition of the key variables. Additional justification for some of the notation used here will be provided in the course of the paper.

$$\nabla \bullet E_C^S = \rho/\varepsilon . \tag{1}$$

$$\nabla \times E_C = 0 . \tag{2}$$

$$\nabla \bullet B = 0 . \tag{3}$$

$$\nabla \times E_I = -\partial B/\partial t . \tag{4}$$

$$\nabla \times B = \mu J_T + \mu\varepsilon\, \partial E_C / \partial t + \mu\varepsilon\, \partial E_I/\partial t . \tag{5}$$

ε is the electrical permittivity and μ is the magnetic permeability. Eq. (1) is Gauss' law for a static, or quasi-static (field fluctuations are transmitted instantaneously) Coulomb field E_C^S generated by a Coulomb charge distribution ρ. It is a consequence of the inverse square law of the Coulomb field of a point charge. Eq. (2) reflects the fact that the general Coulomb fields, including static and dynamic fields, are conservative and can be expressed as the gradient of a scalar Coulomb potential, $E_C = -\nabla\varphi_C$. (In the foregoing, E_C generally refers to dynamic fields.) Equation (3) states that B is solenoidal. Equation (4) is Faraday's law. Given the magnetic vector potential, A, where $B = \nabla \times A$, it follows from Eq. (4) and the assumption that A is the sole source of E_I, that $E_I = -\,\partial A/\partial t$.

Eq. (5) is obtained from Ampere's circuital law. The right hand side is the sum of true currents, J_T and the displacement current, J_D. J_D is generally the sum of contributions from time derivatives of E_C and E_I. Self-consistency requires a solenoidal net current source (sum of J_T and J_D) for B.

In the standard derivation for the wave equations in terms of the potentials, one uses A and φ_C in Eq. (5), to obtain,

$$\nabla \times \nabla \times A = \nabla(\nabla \bullet A) - \nabla^2 A = \mu J_T - \mu\varepsilon\, \partial\nabla\varphi_C / \partial t - \mu\varepsilon\, \partial^2 A / \partial t^2 \tag{6}$$

Rearranging terms,

$$\nabla(\nabla \bullet A + \mu\varepsilon\, \partial\varphi_C / \partial t) = \nabla^2 A - \mu\varepsilon\, \partial^2 A / \partial t^2 + \mu J_T . \tag{7}$$

The two terms in parentheses on the left hand side of Eq. (7) are assumed to be independent and unrelated, with $\nabla \bullet A$ treated as indeterminate. The rationale is that, since $B = \nabla \times A$, any gradient function can be added to A without affecting B. If somehow the left hand side could be set to zero, Eq. (7) would represent an inhomogeneous wave equation with source term, J_T. The conventional approach for obtaining a wave equation from Eq. (7) invokes the electromagnetic gauge which transforms the laboratory system of variables to new variables so that the left hand side of the *transformed version* of Eq. (7) is zero (the Lorenz condition). More specifically, the procedure involves application of the gauge transformation function, ψ, such that

$$A \rightarrow A^{'} + \nabla\psi \tag{8}$$

and

$$\varphi_C \rightarrow \varphi_C^{'} - \partial\psi / \partial t . \tag{9}$$

These transformations allow arbitrary alterations to the variables without affecting the total electric field, E. It is readily seen from Eqs. (8) and (9) that

$$E = E_C + E_I = E_C^{'} + E_I^{'}, \tag{10}$$

where $E_C = -\nabla\varphi_C$, $E_I = -\partial A/\partial t$, $E_C^{'} = -\nabla\varphi_C^{'}$, and $E_I^{'} = -\partial A^{'}/\partial t$.

To obtain the wave equation, first substitute Eqs. (8) and (9) into the left hand side of Eq. (7), and rearrange to give,

$$\nabla(\nabla \bullet A^{'} + \mu\varepsilon\,\partial\varphi_C^{'} / \partial t) = \nabla(\nabla \bullet A + \mu\varepsilon\,\partial\varphi / \partial t - \nabla^2\psi + \mu\varepsilon\,\partial^2\psi / \partial t^2) . \tag{11}$$

One obtains the Lorenz condition,

$$\nabla(\nabla \bullet A^{'} + \mu\varepsilon\,\partial\varphi_C^{'} / \partial t) = 0 , \tag{12}$$

from Eq. (11), by selecting the function ψ such that its values in time and space give zero for the sum of terms in parentheses on right hand side of Eq. (11), i.e.,

$$\nabla^2\psi - \partial^2\psi / \partial t^2 = (\nabla \bullet A + \partial\varphi_C / \partial t) . \tag{13}$$

If we now return to Eq. (7), and apply Eqs. (8) and (9), and the Lorenz condition (Eq. (12)), we obtain,

$$\nabla^2 A^{'} - \mu\varepsilon\,\partial^2 A^{'} / \partial t^2 + \mu J_T = 0 . \tag{14}$$

Eq. (14) is the inhomogeneous wave equation for the transformed vector potential (primed variables).

To complete the task of solving for the total field E, one needs the corresponding scalar wave equation in the transformed variables. This requires the dynamic form of Gauss' law, which is derived from the total electric field,

$$E = E_C' + E_I' . \tag{15}$$

The key step here is to assume Maxwell's Eq. (1) holds for the total field, E. So, inserting Eq. (15) into Maxwell's Eq. (1) gives the familiar result,

$$\nabla \bullet (E_C' + E_I') = \rho/\varepsilon . \tag{16}$$

Note that Eq. (16) must also hold for the unprimed (untransformed) variables since it involves the sum of the two fields.

Expressed in terms of potentials, Eq. (16) becomes,

$$\nabla \bullet (-\nabla\varphi_C' - \partial A' / \partial t) = \rho/\varepsilon . \tag{17}$$

Inserting the Lorenz condition (Eq. (12)) into Eq. (17) gives the scalar wave equation,

$$\nabla^2\varphi_C' - \mu\varepsilon\,\partial^2\varphi_C'/\partial t^2 = -\rho/\varepsilon. \tag{18}$$

The curl and time derivative of the solution for (14) gives B, and E_I', respectively. The gradient of the solution for Eq. (18) gives E_C'. Summing E_C' and E_I' completes the task of obtaining the total electric field E in Eq. (15), which is independent of the choice of gauge function.

The preceding outlines the standard formalism for the elementary derivation of the wave equations. The remainder of this paper addresses questions regarding the validity of these standard results and offers alternatives that correct a variety of errors.

The Missing Maxwell Equation

The first item that we address is the missing Maxwell equation in the standard formalism. Our approach maintains the distinctions between the different physical variables such as E, E_I, and E_C. While this leads to a proliferation of variables, it provides more clarity. In the literature, the same symbol, E, is used for all electric field variables, which obscures the physics and leads to errors, as we will illustrate (see also, the Appendix). The same problem exists with the various components of the vector potential.

The fact that the basic set of Maxwell's equations is incomplete is demonstrated by an error in the standard derivation of the wave equation for the electric field. The conventional approach proceeds as follows (see, for example, Panofsky and Phillips, Chapter 11 [1], Jackson, Chapter 7 [2] , Feynman et al, Chapter 20 [3]): take the curl of both sides of Eq. (4), which gives,

$$\nabla \times \nabla \times E_I = \nabla(\nabla \bullet E_I) - \nabla^2 E_I = -\partial(\nabla \times B)/\partial t . \tag{19}$$

From Eq. (19) and Maxwell's Eq. (5) we have,

$$\nabla(\nabla \bullet E_I) - \nabla^2 E_I = \mu \partial J_T / \partial t - \mu\varepsilon \partial^2 E_C / \partial t^2 - \mu\varepsilon \partial^2 E_I / \partial t^2 . \tag{20}$$

The final step invokes Maxwell's Eq.(1) to justify setting $\nabla \bullet E_I = 0$. Treating the first two terms on the right of Eq. (20) as external source terms gives inhomogeneous wave equation,

.

$$\nabla^2 E_I - \mu\varepsilon \partial^2 E_I / \partial t^2 = \mu J_T + \partial^2 E_C / \partial t^2 . \tag{21}$$

(We derive Eq. (21) later from the vector potential wave equation.) Maxwell's Eq. (1) is usually invoked again at this point [1, 2, and 3] to provide the standard proof, using $\nabla \bullet E_I = 0$, that the vector E_I is transverse to B and to the direction of wave propagation.

Glossing over distinctions among variables is the likely explanation for the improper use of Maxwell's Eq.(1) to set $\nabla \bullet E_I = 0$. The variable E_I refers to an induced field, which is obtained from the vector potential, while Maxwell's Eq. (1) holds only for the static Coulomb field, with the Coulomb charge density as its source term.

It follows from this example that the standard wave equation for E_I has yet to be derived properly. One needs an expression for $\nabla \bullet E_I$ for a valid derivation. Another direct indication that something is missing from the basic set of equations is that the dynamic form of Gauss' law (Eq. (16)) is supposed to give an expression for the source density $\nabla \bullet E_I$ in the presence of a dynamic Coulomb field, while no such expression exists for the static, quasi-static, or charge free cases.

Given that the validity of equation (21) is not in doubt, the missing Maxwell equation must be

$$\nabla \bullet E_I = 0 . \tag{22}$$

More complete justifications for Eq. (22) are provided in the following. We will also show that Eq. (22) does not conflict with the correct form of the dynamic Gauss' law.

Vector Potential

The general expression for the vector potential, A_G, due to a general current density, J_G, is derived, for example, in Panofsky and Phillips [1],

$$A_G = \frac{\mu}{4\pi}\int \frac{J_G(r')}{|r-r'|} dv'. \tag{23}$$

A key point here is that A_G is only defined within an arbitrary function that has a vanishing curl. As shown in Panofsky and Phillips [1], any three dimensional vector is fully characterized by giving its curl and its divergence. Generally, if

$$\nabla \times F = K, \tag{24}$$

so that K is solenoidal, and,

$$\nabla \bullet F = s, \tag{25}$$

it follows that F may be expressed as,

$$F = -\nabla \varphi_F + \nabla \times L, \tag{26}$$

where,

$$\varphi_F = \frac{1}{4\pi}\int \frac{s(r')}{|r-r'|} dv', \tag{27}$$

and,

$$L = \frac{1}{4\pi}\int \frac{K(r')}{|r-r'|} dv'. \tag{28}$$

Consider a general vector potential, A_G where $\nabla \times A_G = B$. An expression for $\nabla \bullet A_G$ (or φ_A) is required for complete characterization of A, so that, generally,

$$A_G = -\nabla \varphi_A + \nabla \times L_A, \tag{29}$$

where

$$L_A = \frac{1}{4\pi}\int \frac{B(r')}{|r-r'|} dv' . \tag{30}$$

Similarly, for E_I, we have $\nabla \times E_I = -\partial B/\partial t$, and, allowing the possibility of a source potential, φ_I, the complete E_I is given by,

$$E_I = -\nabla\varphi_I + \nabla \times L_I , \tag{31}$$

where, invoking Maxwell's Eq. (4), we have

$$L_I = \frac{1}{4\pi}\int \frac{(-\partial B(r')/\partial t)}{|r-r'|} dv' . \tag{32}$$

Maxwell's Eq. (4) (the original Faraday's law) gives only the curl of E_I. The divergence of E_I is also required in order to fully characterize this variable. In other words, the variable E_I is undefined so long as $\nabla \bullet E_I$ remains undefined. So, we have returned here to the case of the missing Maxwell equation (Eq. (22)). If E_I were arbitrary, a broad range of commonplace physics and engineering problems involving induced fields could not be addressed. As discussed in the example of a closed circular loop in the Appendix, approximately a century and a half of experience with applications of the Faraday law indicates that no point sources exist for E_I. Consequently, the required equation for a closed loop of current is $\nabla \bullet E_I = 0$ (Eq. (22)). This equation must be added to the usual statement of Faraday's law (Maxwell's Eq. (4)) for a complete expression of Faraday's law.

As reviewed in the Introduction, the standard formalism treats $\nabla \bullet A_G$ as completely arbitrary. In view of the above discussion of the Faraday law, this approach is no longer acceptable. Consider the net vector potential A. Since $\partial(\nabla \bullet A)/\partial t = \nabla \bullet E_I = 0$, it follows that $\nabla \bullet A$ is restricted to a time independent function. Consequently, when Eq. (22) applies,

$$\partial(\nabla \bullet A)/\partial t = 0 . \tag{33}$$

How is Eq. (33) reconciled with special relativity where the Lorenz condition (Eq. (12)) is generally applied to both A and A'? Again, we will show the answer lies in preserving distinctions among variables.

Alternative Approach to the Wave Equations

Retarded Fields

Among the key equations needed for our purposes are the retarded Coulomb and vector potentials of a moving charge. These potentials deal with the essential features and thus provide a more substantive basis for the development of the wave equations than that used in the standard formalism. A further advantage is that these equations are consistent with the requirements of special relativity. The expressions for the Coulomb and vector potentials given in Feynman et al [3] for a charge moving along the x axis at velocity v are:

$$\varphi_C = \frac{q}{4\pi\varepsilon\sqrt{1-v^2/c^2}\left(\dfrac{(x-vt)^2}{1-v^2/c^2}+y^2+z^2\right)^{1/2}}, \tag{34}$$

and

$$A_x = \frac{(q/c^2)\,v}{4\pi\varepsilon\sqrt{1-v^2/c^2}\left(\dfrac{(x-vt)^2}{1-v^2/c^2}+y^2+z^2\right)^{1/2}};\ A_y = A_z = 0. \tag{35}$$

Recall that these equations reflect the fact that the potentials and vector fields at a point r are given by the moving charge and current located at a point $r'(t')$ that is different from the present source position $r'(t)$ because of the finite speed of light; in these cases $t' = t - \left|r(t)-r'(t)\right|/c$. The term $\sqrt{1-(v/c)^2}$ originates from the Lienard-Weichert expression for the potential and reflects the apparent elongation of the charge in the direction of motion due to retardation effects [3].

For later reference, the concept of retarded fields requires a modification to the general expression (Eq. (23)) for a general vector potential A_G due to a general current source, $J_G(r',t')$. The “retarded solution” version [1,2] of Eq. (23) is:

$$A_G(r,t) = \frac{\mu}{4\pi}\int \frac{[J_G(r',t')]}{\left|r(t)-r'(t)\right|}dv', \tag{36}$$

where the source term in square brackets is evaluated at time $t' < t$.

The Lorenz Condition

Our approach relies upon Eqs. (34) and (35) to obtain basic relationships among the variables. Eqs. (34) and (35) replace the vague notion of "dynamic fields" with precise definitions. It is assumed here that "dynamic fields" is synonymous with retarded fields. The same relationships hold for any charge distributions, in view of the principle of superposition for sums over distributions of charges q_i and distributions of current elements, $q_i v_i$. Note that $A_x = v\varphi_C / c^2$.

Eq. (35) is the vector potential associated with a moving charge and is therefore associated with an element of the true current, J_T. We now label this true current component of the vector potential as A_T, where generally $A_T = A_{TX}i + A_{TY}j + A_{TZ}k$ (i, j, and k are the unit vectors). There are further contributions to the net vector potential, A, from displacement currents, which will be considered when we discuss the vector wave equation. For now, we examine only the relationships between the variables pertaining to true currents. The first item is that the relation between the Coulomb field due to a moving charge and the vector potential due to the current due to that moving charge is now expressed as $A_T = v\varphi_C/c^2$.

Next, if one computes the divergence of A_T, and the time derivative of φ_C, one obtains,

$$\nabla \bullet A_T = \frac{-q\,v\,(x\text{-}vt)}{4\pi\varepsilon\ c^2\left(1\text{-}(v/c)^2\right)^{3/2}\left(\dfrac{(x\text{-}vt)^2}{1\text{-}(v/c)^2}+y^2+z^2\right)^{3/2}}, \tag{37}$$

and

$$\frac{\partial\varphi_C}{\partial t} = \frac{q\,v\,(x\text{-}vt)}{4\pi\varepsilon\ \left(1\text{-}(v/c)^2\right)^{3/2}\left(\dfrac{(x\text{-}vt)^2}{1\text{-}(v/c)^2}+y^2+z^2\right)^{3/2}}. \tag{38}$$

In this case, J_T is a non-solenoidal true current (because it is due to a single moving charge) and $\partial(\nabla \bullet A_T)/\partial t = \nabla \bullet E_I^T \neq 0$ as the reader can verify by taking the time derivative of Eq. (37).

Comparing Eqs. (37) and (38) shows

$$\nabla \bullet A_T + (\partial\varphi_C / \partial t)/c^2 = 0. \tag{39}$$

Eq. (39) closely resembles the Lorenz condition, Eq.(12) which is used to obtain the primed scalar wave equation (Eq. (18)). There are important differences, however: i) unlike the Lorenz condition, which holds for the net primed vector potential, A', Eq. (39) applies to only one component of the net vector potential A, so it does not violate Eq. (33); ii) Eq. (39) is an equation, not a "condition" to be met by adjusting variables; and, iii) Eq. (39) is in covariant form in the unprimed variables, which contrasts with the Lorenz condition. The Lorenz condition does not apply in the unprimed variables, a fact that is generally overlooked in comparisons with results of special relativity.

The fact that an equation that is similar to the Lorenz condition can be obtained in the unprimed variables is well known. For example, Panofsky and Phillips [1] apply the continuity condition, $J_T + \partial\rho / \partial t = 0$ (which refers to a true current in a finite length conductor, with a time varying charge density at the end). They then apply Eq. (36), giving (without subscripts),

$$A_T(r,t) = \frac{\mu}{4\pi}\int \frac{[J_T(r', t')]}{\left|r(t)\text{-}r'(t)\right|} dv' \tag{40}$$

along with an analogous expression for the scalar potential,

$$\varphi_C(r,t) = \frac{1}{4\pi\varepsilon}\int \frac{[\rho(r', t')]}{\left|r(t)\text{-}r'(t)\right|} dv', \tag{41}$$

to show that Eq. (39) is a result of the continuity condition. What is generally overlooked it that the vector potential in this case is the component related to the true current rather than the net vector potential. We use subscripts to preserve the distinctions, which, as we have stressed, makes all the difference.

We obtained Eq. (39) without invoking the continuity condition; however, if one views current in a wire as the flow of a line of charges with drift velocity v over a stationary line of opposite charges, then Eqs. (34) and (35) can be viewed as a similar effect where the dynamic potentials are related to the development of an excess charge at the end of the conductor.

A final point is that some of the more subtle features of relationships among component variables tend to be ignored in the standard approach. Eq. (39) shows, for example, that in the presence of time varying Coulomb fields, every point in space with non-zero $\partial\varphi_C / \partial t$ acts as a point source contribution to A_T (see Eq. (29)). While this component of A_T cannot contribute to a magnetic field B, its time variation does contribute an induced field E_I^T which, in principle, can be detected by the force it produces on charged particles.

The Dynamic Gauss' Law

In the standard approach, the general expression for the divergence of the combined electric fields in the presence of dynamic Coulomb fields is derived as shown for Eqs. (15) and (16). All approaches simply apply the static form of Gauss' law (Maxwell's Eq.(1)) to the sum of the dynamic variables, E_C and E_I, giving the expression,

$$\nabla \bullet (E_C + E_I) = \rho / \varepsilon . \tag{42}$$

No conceptual basis for the new dynamic variables is provided. Despite this, Eq. (40) is generally accepted throughout the physics literature, presumably because it provides the standard gauge transformed scalar wave equation (Eq. (18)). Without it, one cannot obtain the total electric field, E. We now show that Eq. (42) and its equivalent, Eq. (16) are incorrect.

First, assume for the moment that Eq. (40) is valid. The original $\nabla \bullet E_C^S$ (Eq.(1)) and $\nabla \bullet E_I$ (Eq. (22)) are not required to apply in Eq. (42). Instead, the divergence of the *sum* of the two fields at a given point in space is equal to the charge density at that point. This conflicts with the Faraday law: the complete Faraday law must always apply, so Eq. (22), $\nabla \bullet E_I = 0$, always applies. Consequently, Maxwell's Eq. (1) holds for both the static and dynamic Coulomb fields so that there is no need for the concept of dynamic fields if Eq. (42) is correct. The source of the difficulty is that the standard assumption that Maxwell's Eq. (1) applies to the sum of E_C and E_I is simply invalid.

A proper derivation of the dynamic form of Gauss' law is available from the retarded field expressions for φ_C and A_T (Eqs. (34) and (35)). For the sake of brevity we do not include the rather lengthy equations that result from the straightforward derivatives, but give only a summary of the results, which the reader can readily verify. If one obtains E_C from $-\nabla\varphi_C$ and E_I^T from $-\partial A_T / \partial t$ outside the singularity, one finds that $\nabla \bullet E_C \neq 0$ and $\nabla \bullet E_I^T \neq 0$.

On the other hand, if one now uses Eq. (34) and (35) to compute the divergence of the *sum* of these two vectors outside the singularity, one obtains,

$$\nabla \bullet (E_C + E_I^T) = 0 . \tag{43}$$

In order to include the singularity at the location of the charge, q, at (x-vt), y, z, we use the fact that the correction term $\sqrt{1-(v/c)^2}$ is only valid distances larger than the size of the charge [3]. Retardation effects cannot exist in the immediate vicinity of the charge, so the fields there follow the static inverse square law for E_C^S. The result is a correct expression for the dynamic form of Gauss' law,

$$\nabla \bullet (E_C + E_I^T) = q/\varepsilon\,. \tag{44}$$

Applying the principle of superposition for a distribution of charges as in the static case gives the more general expression,

$$\nabla \bullet (E_C + E_I^T) = \rho/\varepsilon\,. \tag{45}$$

This result reveals a previously unrecognized feature of the dynamic Gauss' law, namely, that it holds for the dynamic Coulomb field and for only one component of the induced field, E_I^T. There are no contributions from the induced fields generated by the two displacement currents.

Returning to the topic of variable labels, note that Maxwell's Eq. (2) holds for general Coulomb fields. Thus, it applies to both the static E_C^S and to the dynamic E_C described here.

The Scalar Wave Equation

The correct form of the dynamic Gauss' law reveals a flaw in the gauge approach. The key step in obtaining the scalar wave equation relies on the incorrect form of Gauss' law, (Eq. (16)). We now see that there is no such difficulty with the unprimed variables and the correct version of the dynamic Gauss' law.

Rewriting Eq. (45),

$$-\nabla^2 \varphi_C - \partial(\nabla \bullet A_T)/\partial t = \rho/\varepsilon\,. \tag{46}$$

Applying Eq. (39) to Eq. (46) gives,

$$\nabla^2 \varphi_C - \partial^2 \varphi_C / \partial t^2 / c^2 = -\rho/\varepsilon\,, \tag{47}$$

which is the scalar wave equation in unprimed variables. Eq. (47) is independent of the vector wave equation. This contrasts with the results of the gauge approach, where the primed scalar and vector wave equations are actually meaningless by themselves. (For example, even in cases where actual Coulomb fields are absent, the gauge approach requires that one consider an imagined Coulomb field.) Both primed equations are coupled because only the sum of the primed electric fields is accessible to laboratory testing. A further advantage of Eq. (47) and the vector wave equation (derived in the next section) over that obtained from the gauge approach is that they can legitimately be discussed in terms of the unprimed variables that can be measured directly or compared with results from special relativity.

The Vector Wave Equation

We begin with Eq. (6), which is Maxwell's Eq. (5) in terms of the vector potential,

$$\nabla\times\nabla\times A=\nabla(\nabla\bullet A)-\nabla^2 A=\mu J_T+\mu\varepsilon\,\partial E_C/\partial t+\mu\varepsilon\,\partial E_I/\partial t\,. \tag{6}$$

Eq. (6) shows there are three source terms for the net vector potential, A. To underscore the point that A is actually the sum of components, we consider each separately. The component, A_T, due to the true current has already been discussed. The explicit retarded field expression for the vector potential component due to the Coulomb displacement current can be written as

$$A_C(r,t)=\frac{\mu}{4\pi}\int\frac{[\varepsilon\,\partial E_C(r',t')/\partial t]}{|r(t)\text{-}r'(t)|}dv'\,, \tag{48}$$

and the third component of A, due to the displacement current from the total induced field, E_I, can be written as

$$A_I(r,t)=\frac{\mu}{4\pi}\int\frac{[\varepsilon\,\partial E_I(r',t')/\partial t]}{|r(t)\text{-}r'(t)|}dv'\,. \tag{49}$$

The net vector potential is the sum of these three terms,

$$A\text{=}A_T+A_I+A_C\,. \tag{50}$$

Since the source current for net A is solenoidal, the general criterion for $\partial(\nabla\bullet A)/\partial t=0$ is satisfied. Taking the time derivative of both sides of Eq. (6), and employing $\partial(\nabla\bullet A)/\partial t=0$, gives the wave equation for the induced field, E_I,

$$\nabla^2 E_I\ \text{-}\ \mu\varepsilon\,\partial^2 E_I/\partial t^2=\mu\,\partial J_T/\partial t+\mu\varepsilon\,\partial^2 E_C/\partial t^2\,. \tag{51}$$

Thus, with the present approach, one obtains exactly the same wave equation for E_I as that obtained directly (Eq. (21)) from the basic set of Maxwell's equations. As with the scalar equation, the vector wave equation holds for the unprimed variables. This consistency check is not possible with the gauge approach since, by definition, the meanings of the primed variables differ from those of the unprimed variables.

For completeness, we note that the corresponding wave equation for B is obtained from the curl of both sides of Eq. (6). Using Maxwell's Eq. (2) and the fact that $\nabla\times\nabla(\nabla\bullet A)\text{=}0$ gives,

$$\nabla^2 B - \mu\varepsilon \partial^2 B / \partial t^2 = \mu \partial(\nabla \times J_T) / \partial t, \tag{52}$$

which is identical to that obtained directly from Maxwell's Eqs. (2), (3), and (4).

Another consistency check is that the solenoidal source terms in both Eqs. (51) and (52) generate the solenoidal vector fields, B and E_I. The same is not true of Eq. (6) since both sides of Eq. (6) are solenoidal, regardless of $\nabla \bullet A$. On that topic, we have shown that the basic wave equations (i.e., Eqs. (49), (50), and the scalar wave equation (Eq. (47))) can be derived from the potentials without invoking an electromagnetic gauge. The correct wave equations are obtained for any time independent function for $\nabla \bullet A$. The reason can be seen from Eq. (29). A constant $\nabla \bullet A$ generates neither a magnetic field nor an electric field, so it is physically undetectable. It is physically meaningless. Furthermore, there can be no physical justification for the existence of such a term where solenoidal current sources generate continuous flow lines for A. So, the concept of a source term for A is physically meaningless, which requires $\nabla \bullet A = 0$. Equation (6) can then be written as,

$$\nabla^2 A - \partial^2 A / \partial t^2 = -\mu J_T - \mu\varepsilon \partial E_C / \partial t. \tag{53}$$

As a final point, the present results reflect a general underlying physical principle: solenoidal current sources generate solenoidal fields so the statement that there are no monopoles for B, may be extended to include E_I and A.

Reconciling with Special Relativity

The direct incorporation of gauged (or transformed) variables into the equations for special relativity is not correct because one cannot equate primed and unprimed variables. We already noted that $A_T = v\varphi_C$, where is the component due to the true current, and that the unprimed analog to the Lorenz condition is also expressed in terms of A_T. Bearing this in mind, if one re-examines the primed vector potential wave equation (Eq. (14)) it is clear that it also applies only to A_T. The displacement contribution from the Coulomb field is removed in the gauge transformation process leaving a non-solenoidal source term and a wave equation for a non-solenoidal vector field. In other words, the primed net vector potential is the same as the component due to the true current, $A' = A_T$. Thus, the main result of the standard gauge in electromagnetism is a wave equation that is applicable to only A_T. Since the standard covariant expressions are based on the primed vector fields, it is clear that all such expressions actually refer to the A_T component, rather than the net A. Hence, simply changing the label of the A vectors in the special relativity equations to A_T corrects the misleading equations and removes the conflict between the present results and the requirements of special relativity. Reiterating, the present results show that Maxwell's equations are consistent with $\nabla \bullet A = 0$, while the relativistic equations are consistent with $\nabla \bullet A_T + (\partial \varphi_C / \partial t) / c^2 = 0$.

A similar situation exists in the quantum mechanical treatment of electromagnetism. Feynman and Hibbs [4], for example, discuss the fact that a quantum mechanical formalism that incorporates the Lorenz condition is considerably more cumbersome than that using $\nabla \bullet A=0$. So $\nabla \bullet A=0$ is generally assumed to apply. Again, the conflict between this requirement and special relativity is resolved by recognizing that the relativistic equations refer to A_T rather than A.

Summary

We have shown that much of the standard formalism of electromagnetism cannot survive the application of precise definitions to its variables. The key to addressing this problem is the inclusion of an additional equation to the basic set of Maxwell's equations. This additional equation is required to complete the statement of Faraday's law. With a complete set of basic equations, all the essential relationships needed for valid derivations of the wave equations can be obtained without an electromagnetic gauge.

References

1. Panofsky, W.K.H., and Phillips, M., *Classical Electricity and Magnetism* (2nd Ed. Addison-Wesley, Reading MA, 1962).

2. Jackson, J.D., *Classical Electrodynamics* (John Wiley and Sons, NY, 1962).

3. Feynman, R., Leighton, R., and Sands, M., *The Feynman Lectures in Physics vol II* (Addison-Wesley, Reading, MA, 1964).

4. Feynman, R., and Hibbs, A. R., *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).

5. Cote, P.J., Johnson, M.A., Truszkowska, K, and Vottis, P., J. Phys D Appl. Phys. **40** 274-283 (2007).

Appendix

In the present paper we attribute errors in the standard derivations of wave equations to the common practice of neglecting distinctions among key field variables. Reference [5], describes similar errors in a variety of textbook analyses related to applications of the Faraday law. This Appendix provides a simple illustrative example of such errors using a circular shell where no Coulomb fields exist. This example also serves to support several of the central points made in the main text.

Consider a thin, conducting circular cylindrical shell of 1 meter diameter, with a loop resistance of 1 ohm, enclosing a uniform magnetic field which is increasing linearly in time so that the rate of change of flux (emf) is 1 volt (Figure 1). (The shell is thin enough that skin effect can be neglected.) What is the potential difference between points A and B?

A common approach is to apply the Faraday law in the following manner: Using the fact that the line integral around any closed path is equal to the negative of the flux enclosed, one conventionally assumes: i) that no fields exist inside a good conductor and ii) any convenient path can be selected to give the field between points A and B. The simplest choice is a closed path that passes along a zero field circumferential segment from A to B inside the shell and exits the shell at B and follows the diameter back to A to close the loop. This loop encompasses half the area, so the non-zero field between A and B is 0.5 volts/meter and the potential between A and B is 0.5 volts. If one considers the symmetry of the setup, however, this must be incorrect.

The root of the problem is ignoring the distinctions between Coulomb fields and induced fields. Potentials are obtained from line integrals of Coulomb fields while emf's are obtained from line integrals of induced fields. Furthermore, paths cannot be chosen arbitrarily: the induced fields, E_I, actually form closed loops that are concentric with the conducting shell, (and inside the shell itself). Because perfect circular symmetry is assumed, there can be no induced charges, so there are no Coulomb fields; hence, there are no potential differences. So, there are no fields directed along the diameter between points A and B. Furthermore, the fields inside the shell are not zero, so a 1 amp current will be generated. The final point is that the physics of the problem cannot described by the sum of E_I and E_C as required in the gauge approach.

This example also relates to requirement for a full characterization of E_I and the possible presence of a source term for E_I in the absence of a Coulomb field. A source term implies a potential, φ_I (See Eqs. (24) through (28)). Thus, the line integral of E_I between any two points on the circular shell would include a contribution from the gradient of φ_I. The result would be an accumulation of charges at different points along the shell, which, in turn, would produce detectable potential differences. (A concrete example, consistent with the circular symmetry of the present example, is a source term for E_I along the central axis of the shell. This would produce a potential difference between the inner and outer surfaces of the shell.) Such sources produce no effect on the total current, however, since the line integral of a gradient of φ_I around a closed circuit is zero. The presence of a source term would yield deviations between predicted and measured potentials in a wide range of electromagnetic devices. In this case, one can fairly assume that the absence of evidence of a source term is evidence of its absence, so Eq. (22) must apply for closed current loops.

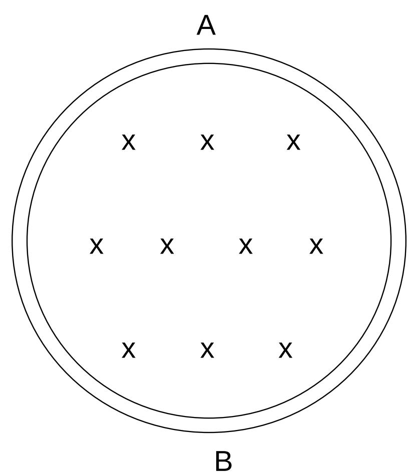

Figure 1: Conducting shell enclosing a uniform, time-varying magnetic field.